

# SYSTEMS AND METHODS FOR PREDICTING USAGE OF A WEB SITE USING PROXIMAL CUES

## INCORPORATION BY REFERENCE

5       **[0001]** This Application incorporates by reference Attorney Docket No. D/A0045, entitled "SYSTEM AND METHOD FOR PREDICTING WEB USER FLOW BY DETERMINING ASSOCIATION STRENGTH OF HYPERMEDIA LINKS", by P. Pirolli et al., filed March 31, 2000 as US. Application Serial No. 09/540,976; and Attorney Docket No. D/99794, entitled "SYSTEM AND METHOD  
10 FOR INFERRING USER INFORMATION NEED IN A HYPERMEDIA LINKED DOCUMENT COLLECTION", by E. Chi et al., filed March 31, 2000 as U.S. Application Serial No. 09/540063, each in its entirety.

**[0002]** The U.S. Government has a paid-up license in this invention and the right in limited circumstances to require the patent owner to license others on  
15 reasonable terms as provided for by the terms of Contract No. N00014-96-C-0097 awarded by the Office of Naval Research.

## BACKGROUND OF THE INVENTION

### 1. Field of Invention

**[0003]** This invention relates to predicting the usage patterns of document  
20 collections given information about a user's information needs.

### 2. Description of Related Art

**[0004]** Increasingly, the World Wide Web has become the information delivery mechanism of choice for both corporations and individuals users. The ubiquity of World Wide Web browsers and the push by many corporations to adopt  
25 common off the shelf technology (COTS) have all helped the World Wide Web become a required delivery option for most information systems.

**[0005]** However, although information sources are now more likely to be available to their intended audience through the World Wide Web, the access to relevant information is still limited by a user's ability to navigate the World Wide  
30 Web and the destination web site and to actively accumulate the required information. Many sites use different methods or models of site design to present the information. For example, a web site designer of a county government tax assessors office site may assume any query will be related to county tax assessment. In contrast, the web site

designer for an online department store needs to provide a user with access to product information ranging from toasters to jewelry. The web site designer of an internal corporate information site may need to provide access to corporate tax information, real estate holdings, business permits and/or health and safety records. Providing an intuitive interface to facilitate user access to information repositories including web sites is therefore increasingly important for businesses and consumers.

[0006] Accordingly, Web site designers, information system managers, and researchers are constantly developing new tools to gain understanding into the paths that users follow to obtain the information they need. For example, Web site designers, researchers and web site banner advertisers seeking to place information on the most relevant web site have used a variety of techniques to analyze web log files. Web log files contain information concerning which web page referred the user to the site as well as which web pages were visited within the site. Information concerning the user's IP address and browser type is also frequently saved for review in the web log file.

[0007] Tools such as Insight from Accrue Corporation, Astra Site Manager from Mercury Interactive and WebCriteria's Site Analysis, Task Analysis and MAX products allow Web site developers to analyze general statistics about a web site. For example, WebCriteria's Site Analysis product provides descriptive statistics accumulated through the use of the MAX software agent product. The MAX software agent traverses the web site to derive usability metrics from simulated browsing. However, the simulated browsing merely provides a random walk of a web site. Simulated browsing based on a random walk assumes the user's navigational choices at any juncture are random and simply ignores the presence of informational cues on each page and surrounding each link. However, in the actual use of the site, informational cues influence a user's decision as to whether one path through the Web site is chosen over another path.

[0008] In Chakrabarti et al. and Silva et al, (Chakrabarti, S., B. Dom, P. Raghavan, S. Rajagopalan, D. Gibson and J. Kleinberg. Automatic Resource Compilation by Analyzing Hyperlink Structure and Associated Text. In Proc. of the 7<sup>th</sup> International World Wide Web Conference (WWW7), pp. 65-74, Brisbane, Australia, 1998, and Silva, I., B. Ribeiro-Neto, P. Calado, E Moura., N. Ziviani,

Link-based and Content-Based Evidential Information in a Belief Network Model. In Proc. of the 21<sup>st</sup> ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 96-103, Athens, Greece 2000), a combination of keywords and links is used to determine a ranking weight for retrieval results. However,

5 Chakrabarti and Silva make no attempt to predict the usage of a web site based on a virtual user's information needs.

[0009] Instead, these systems merely describe how users have traversed the web site in the past. These systems fail to provide web site designers an objective prediction, useful in describing how the changes to a document or web page affect the way a user with a specific information need will traverse the site. Co-pending Application, Attorney Docket No. D/A0045, entitled "SYSTEM AND METHOD FOR PREDICTING WEB USER FLOW BY DETERMINING ASSOCIATION STRENGTH OF HYPERMEDIA LINKS", by P. Pirolli et al., filed March 31, 2000, and filed as US. Application Serial No. 09/540,976, incorporated in its entirety, predicts a user's traversal of the links in a document collection or web site using a computation based on the presence of information in the linked to or distal page. However, distal information is information which by definition has not yet been seen by the user. Accordingly, analysis of distal information cannot reflect an objective indication of the decisions made by an actual first time user of the document collection or web site, as the user encounters the navigational choices in the current or proximal document or web page. Instead, document collection or web site usability analysis requires including some measure of how the user's experience is affected by information cues in the proximal or current document or web page.

#### SUMMARY OF THE INVENTION

25 [0010] Accordingly, systems and methods for predicting usage of a document collection or web site, using proximal cues would be advantageous. In various exemplary embodiments according to this invention, a virtual or objective user's predicted usage of a document collection or web site using proximal cues is determined.

30 [0011] Various other exemplary embodiments according to this invention provide for identifying and providing proximal information cues for image links in the document collection or web site.

## BRIEF DESCRIPTION OF THE DRAWINGS

[0012] Various exemplary embodiments of this invention will be described in detail, with references to the following figures, wherein:

Fig. 1 is a block diagram showing an exemplary system for predicting the usage of the a web site according to this invention;

Fig. 2 is a flowchart of an exemplary method for predicting the usage of a web site according to this invention;

Fig. 3 is a flowchart of an expanded view of an exemplary method for determining proximal scent according to this invention;

Fig. 4 is a flowchart of an expanded view of an exemplary method for determining distal scent according to this invention;

Fig. 5 is an expanded view of an exemplary method of spreading activation according to this invention;

Fig. 6 is an exemplary graph showing a set of documents to be analyzed according to this invention;

Fig. 7 is an exemplary topology matrix according to this invention; and

Fig. 8 is an exemplary proximal cue word matrix according to this invention.

## DETAILED DESCRIPTION OF EXEMPLARY EMBODIMENTS

[0013] Fig. 1 shows an exemplary block diagram of a first exemplary embodiment of a system for predicting the usage of a web site 100 according to this invention. The system for predicting usage of a web site 100 includes a controller 15, a memory 20, a topology determining circuit 25, a proximal cue determining circuit 30, a word/document frequency determining circuit 35, a keyword matrix determining circuit 40, a non-zero topology element determining circuit 45, a proximal scent determining circuit 50, a normalizing circuit 55, a proximal and distal scent combining circuit 60, a spreading activation circuit 65, a scent matrix determining circuit 75, a keyword x weighted word/document determining circuit 70, a weighted word/document matrix determining circuit 80, a distal scent determining circuit 85 and an input/output circuit 10 for connecting to communications link 110. A web server 95 provides access to a document collection or web site pages 90 over communications link 110.

[0014] The input/output circuit 10 retrieves the document or web page of a web site 90 over the communication link 110 from web server 95 and stores the document or page in memory 20. The exemplary embodiment of this invention describes the use of a web server 95. However, it will be apparent that the system for predicting usage of a web site using proximal cues 100 according to this invention may be practiced using any system or method of providing access to document collections, web sites or any other type of content portions in which links, references and/or connection between the documents, web pages or content portions are provided.

[0015] As the input/output circuit 10 retrieves each document or web page in the site to be analyzed, the topology determining circuit 25 builds an exemplary topology matrix describing the links or connections between each document or web page retrieved. For purposes of description, the exemplary site shows a limited number of documents or web pages and associated links. However, any number of documents or web pages may be processed using the systems and methods of this invention.

[0016] As the input/output circuit 10 retrieves each document page of the document collection or web site 90, the topology of the document or web site is determined using the topology determining circuit 25. Each link or connection on a document or web page is then analyzed by the proximal cue analyzing circuit 30. The proximal cue analyzing circuit 30 identifies information cues that convey information to the virtual user. For example, the text associated with a link or connection may provide proximal cues as the nature of the information linked to. The proximal cue analyzing circuit 30 then breaks the link or connection down into constituent words. The proximal cue words may also include portions of the text surrounding the link. If a link or connection is in typical URL form, the punctuation and '/' characters associated with the URL are used to define word boundaries. For example, if the link or URL "http://www.xerox.com/products/support/index.html" were processed, the words http, www, xerox, com, products, support and index would all be added as proximal cue words to the keyword matrix for the relevant link. In various alternative embodiments, additional weighting factors may be added to the proximal cue word information such as the location of the link within the document or web page. For

example, a weighting that accords higher importance to cues words appearing higher up in the structure of a document or web page may also be used to provide weighting information.

[0017] If the link is an image link such as a GIF, JPEG, PNG, BMP or any other image file type, the number of proximal cue words derivable from the link may be limited or non-existent. In addition, determining proximal cue words for image links is a difficult problem. If the image is analyzed, a large increase in processing time and resources will result. Also, determining the meaning or semantics of an image based on image analysis is an extremely complicated and error prone process. However, as web site developers attempt to improve the usability of their web sites, the use of image links has increased. Accordingly, it is important to include image links when determining the proximal cues a virtual or objective user responds to.

[0018] If the controller 15 determines that the stored document or web page is an image link, then the distal document or web page that is linked to the image is retrieved. In the exemplary embodiment of this invention, the proximal cue word information from the link and/or text surrounding the link are analyzed by the proximal cue calculating circuit 30. The linked to or distal document or web page may also be analyzed for proximal cue words by the proximal cue calculating circuit 30. In this way, the linked to or distal document or web page serves as a proxy to supplement or replace proximal cue word information insufficient or unavailable on the current or proximal page for the image link. In various alternative embodiments, the linked to or distal information used may include the distal document title and the document text, either alone or in combination with proximal cue information such as the cue words from the text surrounding the image link. Entries reflecting the presence of the proximal cues are then stored in the keyword matrix data structure stored in memory 20.

[0019] Once the topology determining circuit 25 has determined all the topology information for a document collection or web site, the information is stored in the exemplary topology matrix data structure stored in memory 20. The word/document frequency determining circuit 35 determines the word / document frequency of each of the relevant words in the document collection or set of web pages making up the web site 90. The weighted word/document determining circuit

80 then determines the weighted term frequency, inverse document frequency for each of the words in the document collection or web site.

5       **[0020]** The non-zero topology element determining circuit 45 then analyzes the exemplary topology matrix data structure stored in memory 20. Each of the non-zero elements of the topology matrix stored in memory 20 are then identified.

10       **[0021]** For each of the identified non-zero topology elements in memory 20, the controller 15 identifies the relevant link/document from the topology matrix. The link/document identification information is then passed to the keyword matrix determining circuit 40. The keyword matrix determining circuit 40 analyzes the proximal matrix stored in memory 20 for entries corresponding to the link/document identification information. The non-zero entries in the proximal cue word matrix entry for the link/document indicate relevant words associated with the link.

15       **[0022]** For each non-zero entry specified in the proximal cue word vector, proximal scent determining circuit 50 then determines the proximal scent by determining the similarity between each entry in the proximal cue word matrix stored in memory 20 and the user information need vector previously stored in memory 20. In the exemplary embodiments according to this invention, the similarity is determined by multiplying the proximal cue word matrix stored in memory 20 by the weighted word/document frequency matrix stored in memory 20 using proximal cue word x weighted word/document determining circuit 70 and multiplying the result by the user information need vector previously stored in memory 20. However, in various other exemplary embodiments according to this invention, any known or later developed technique of determining similarity may be employed. For example, a determination of the cosine of the angle between the vectors may be used.

25       **[0023]** The distal scent determining circuit 85 determines the distal scent matrix using information from the connected to or distal document or web pages. The connected to or distal documents or web pages provide the cue information to determine the distal scent matrix.

30       **[0024]** The proximal and distal scent combining circuit 60 is then activated to substitute the previously determined distal scent matrix entries into the proximal scent matrix when the distal scent matrix entry is non-zero and the proximal scent entry is zero. The new matrix created is called the scent matrix. In this way, distal

information is provided as a substitute when there is insufficient proximal information. It will be apparent that other techniques of determining the proportions of proximal and distal scent to be combined are within the scope of this invention.

For example, the relationship:

5

$$\text{Scent} = \text{ALPHA} * (\text{Proximal\_Scent}) + \text{BETA} * (\text{Distal\_Scent}) \quad (1)$$

might be used.

10 [0025] The normalizing circuit 55 is then activated to create a normalized scent matrix in which the columns of the scent matrix sum to one. This reflects that the sum of all probabilities for a user navigation choice at that particular point is one. Each entry reflects the likelihood that a user with the specified information need will choose the associated path.

15 [0026] The controller 15 then selects the initial or starting page vector E previously stored in memory 20. The initial or starting page vector reflects the first document or page the virtual or objective user selects in traversing the document collection or web site. The first document or page may reflect any document or page within the document collection or web site. The spreading activation circuit 65 is then initialized with the number of iterations to be run. After the specified number of 20 iterations, the matrix is analyzed. The matrix entries reflect the likelihood the virtual or objective user will arrive at the indicated location.

[0027] As a web site designer changes the document collection or web site, an objective indication of the document collection or web site usability can be generated. In this way the web site designer can interactively adjust the design of a 25 web site and develop better techniques to deliver the desired information to the target audience.

[0028] Fig. 2 shows a flowchart of an exemplary method for predicting the usage of a web site according to this invention. The process starts at step S10 and immediately continues to step S15 where the user's information needs are entered in 30 the form of the query. A query vector representing the words appearing in the query is generated. Control then continues to step S20. In step S20, the topology matrix for the document collection or web site is determined. In the exemplary embodiment, the



topology matrix encodes information describing the links that exist between documents or web pages to be analyzed. The exemplary embodiment shows a matrix encoding the information. However, it will be apparent that any type of data structure such as matrices, adjacency lists or any other known or later developed data structure capable of indicating the relationship between entries may be used in the practice of this invention.

[0029] In step S30, the first non-zero link in the topology matrix is selected. The topology matrix encodes information describing which documents, or web pages contain links to other documents or web pages. Therefore, documents containing relevant links are easily identified. Control then continues to step S40 where a determination is made whether the link structure is an image link.

[0030] Image links present a difficult semantic analysis problem since complex image analysis is necessary to determine the subject of the image. The proximal cue words surrounding a link frequently provide some cue words. However, if no proximal cue words can be determined then cue words from the distal document can be used. If it is determined in step S40 that the link structure is an image link, control continues to step S50. If it is determined that the link structure is not an image link, control continues to step S35.

[0031] In step S35, the link indicated by the non-zero topology matrix entry is analyzed for proximal cue words. Proximal cue words include for example, the text of the link structure. For links including characters such as '/' and '.', the characters may be used as word boundary markers and each portion of the text added to the list of proximal cue words. Text surrounding the link structure may also be added, the title of the proximal page as well as features such as where the link is found within the document. It will be apparent that any feature of the proximal document or web page may be used to provide proximal cue words in the practice of this invention. Control then continues to step S60.

[0032] If the image determination step S40 indicates that the link structure is an image, control continues to step S50. In step S50, the cue words on the distal document or web page may be used to provide additional cue words for the image link structure. In the exemplary embodiment of this invention, proximal cue words for the link are determined from the link, the text surrounding the link, the title of the distal

document and/or the text of the distal document, either alone or in combination. If the distal document or web page cue words are insufficient, the image filename used in naming the image link structure can also be used to provide cue words. Control then continues to step S60.

5           **[0033]** The proximal cue words are then added to the proximal cue word matrix K in step S60. The proximal cue word matrix K stores information about which words are associated with which links. Each link in the keyword matrix represented by a row entry is associated with a respective proximal cue word. Each link in the set of links has an entry associated with each respective cue word.

10           **[0034]** Control then continues to step S70 where the word/document frequency matrix is determined W. The word/document frequency matrix indicates how frequently a word as indicated by a column entry appears in the document specified by the row entry. The weighted word/document frequency is then determined using term frequency, inverse document frequency, term frequency, log of  
15 term frequency,  $(1 + \log_{10})$  of term frequency or any other known or later developed technique of weighting. Control continues to step S80 where the proximal scent matrix is determined.

**[0035]** An expanded view of the determination of the proximal scent matrix is discussed in relation to Fig. 3. After the proximal scent matrix is determined in step  
20 S80 control continues to step S90 where the distal scent matrix is determined. An expanded view of the determination of the distal scent matrix is discussed in relation to Fig. 4. Control then continues to step S100 where the proximal and distal scent matrices may be combined. In the exemplary embodiment of this invention, when a distal scent matrix entry is not zero and the proximal scent matrix is zero, the distal  
25 scent matrix entry may be substituted into the proximal scent matrix entry. In this way, the distal scent matrix provides information which is not available through the proximal scent matrix. For example, the combined scent matrix may represent any combination of the distal and proximal scent matrices as shown in equation 2. Control then continues to step S110.

30

$$\text{Scent} = a * \text{Proximal\_Scent} + b * \text{Distal\_Scent}. \quad (2)$$

[0036] In step S110 a starting page for spreading activation is selected. The starting page is represented by a vector E. For example,  $E = \{1\ 0\ 0\ 0\ 0\ 0\}$  represents the starting page P0 in the exemplary document collection or web site of Fig. 6.

Control then continues to step S120 and the spreading activation iteration begins.

- 5 [0037] Spreading activation is applied or iterated 'n' times in step S120. In each iteration of spreading activation, the parameter alpha indicates the percentage of users who will follow the through to the next link. Accordingly, each iteration of the spreading activation models the probability that a user with information needs represented by the query vector Q, starting at page E will traverse the indicated pages.
- 10 In this way a prediction of the usage of a web site by a user having the indicated information needs and starting point is obtained. When the spreading activation has iterated 'n' times, control continues to step S130 and the process ends.

- 15 [0038] Fig. 3 shows an expanded view of the exemplary method for calculating proximal scent according to this invention. The process starts at step S200 and continues immediately to step S210.

[0039] In step S210, the first entry in the topology matrix T is selected and control continues to step S220. In step S220, a determination is made as to whether any topology entries remain to be analyzed.

- 20 [0040] If it is determined that no further link entries remain to be analyzed, control is transferred to step S290 where the proximal scent matrix is normalized. Normalization of the matrix is achieved by performing elementary column operations on the matrix columns to ensure the columns sum to one. Control then continues to step S295 and control is returned to the calling step S80 of Fig. 2.

- 25 [0041] If the determination is made that further entries remain in the topology matrix, control continues to step S230. In step S230 a determination is made whether the topology matrix entry contains a zero value. A zero value in the topology matrix indicates that no link exists. If the determination indicates the topology entry is zero, control is transferred to step S250, otherwise control is transferred to step S240. In the exemplary embodiment of this invention, a zero value indicates no link exists. However, it will be apparent that any value, such as a negative number may be
- 30 used that indicates no link exists in the topology matrix.

[0042] If control is transferred to step S250, the proximal scent matrix entry corresponding to the topology entry is set equal to zero since no information scent can be transferred along a non-existent link. Control then continues to step S280.

[0043] If the determination in step S230 indicates that the topology entry is not zero, then control continues to step S240. In step S240, the proximal cue word vector associated with the link is determined from proximal cue word matrix K. The proximal cue word vector stores information for all the proximal cue words associated with a given link. Control then continues to step S260 where each proximal cue word in the proximal cue word vector is multiplied by the weighted word/document frequency matrix to obtain a second weighted proximal cue word vector. Control then continues to step S270 where the similarity of the second weighted proximal cue word vector to the initial query vector is determined. As discussed above, the initial query vector indicates the user's information needs. In the exemplary embodiment, similarity is multiplying the second weighted proximal cue word vector by the initial query vector. However, it will be apparent that in various other exemplary embodiments, similarity may be determined by for example, using the cosine between vectors, or any other known or later developed technique of determining similarity. Control then continues to step S280.

[0044] In step S280, the next entry in the topology matrix is selected. Control then continues to step S220 where the process repeats until a determination is made in step S220 that no topology link entries remain to be analyzed and control is transferred to step S290 where the proximal scent matrix is normalized. Normalization of the matrix is achieved by performing elementary column operations on the matrix columns to ensure the columns sum to one. Control then continues to step S295 and control is returned to the calling step S80 of Fig. 2.

[0045] Fig. 4 shows an expanded view of an exemplary method of determining distal scent according to this invention. The process begins at step S300 where control is immediately transferred to step S310.

[0046] In step S310, a distal scent matrix is determined. Cue words are determined based on the information content of the distal or connected to document or web page as further discussed in Co-pending Application Serial No. 09/540,976 incorporated above. Control then continues to step S320.

[0047] In step S320 the distal scent matrix is normalized so that the sum of the columns in the matrix equals one. This indicates that the probabilities of the respective path transitions through the collection of documents or web pages. Control then continues to step S330 where the process is returned to the calling step S90 of Fig. 2.

[0048] Fig. 5 shows an expanded view of an exemplary method of spreading activation according to this invention. The process begins at step S400 and control is immediately transferred to step S410.

[0049] In step S410, the initial starting document or web page is stored as the vector E. The initial vector is defined as the first iteration of the spreading activation network. Control then continues to step S420.

[0050] In step S420, the alpha parameter is set. The alpha parameter signifies the percentage of users who click through or traverse each link. Control then continues to step S430 where the number of iterations of spreading activation is entered and the activation counter is set to one. Control then continues to step S440.

[0051] In step S440, a determination is made as to whether the count is greater than or equal to the number of iterations selected in step S430. If the count is less than the number of iterations selected, control jumps to step S445. In step S445, a new activation is determined by applying the following formula (2).

$$A(\text{COUNT}) = \text{ALPHA} * (\text{Scent\_Matrix}) * A(\text{COUNT} - 1) + E \quad (3)$$

Control then continues to step S450 where the count is incremented and the process repeats until the determination in step S440 indicates the count value is greater than or equal to 'n'. If the count is determined to be greater than or equal to 'n' then control jumps to step S460 and the process ends.

[0052] Fig. 6 shows an exemplary graph showing the connecting links between an exemplary set of documents or web pages.

[0053] Fig. 7 shows an exemplary topology matrix T according to this invention. The topology matrix stores information about the connections between the documents or web pages. Non-zero entries in the topology matrix indicate links for which proximal and distal scent matrix entries will be calculated.

[0054] Fig. 8 shows an exemplary proximal cue word matrix K according to this invention. Each word appearing in the documents or web pages is analyzed according to term frequency/ inverse document frequency criteria. Each weighted word appears as a column entry in the keyword matrix. The presence of a non-zero entry in the matrix indicates the proximal cue words associated with the link specified by the row index. Similarly, specific proximal cue words for a given link are identified by corresponding non-zero entries in the columns of the row associated with the link.

[0055] For example, row 0 indicates that the proximal cue words 'SUN' and 'HOME' appear in document P0. Similarly row 1 indicates that 'JAVA' and 'HOME' appear in document P1. Entries in row 2 indicate that the words 'JAVA' and 'SUPPORT' appear in document P2. The non-zero entries in row 3 indicate that 'JAVA' and 'API' appear in document P3. Non-zero row entries in row 4 indicate that 'HOME' and 'PETE'S' appear in document P4. Non-zero entries in row 5 indicate that 'JAVA' and 'COFFEE' appear in document P5. Finally, the non-zero entry in row 6 indicates that the document P6 contains 'TEA'. In this way a keyword matrix is formed.

[0056] In the various exemplary embodiments outlined above, the system for predicting the usage of the a web site 100 can be implemented using a programmed general purpose computer. However, the system for predicting the usage of the a web site 100 can also be implemented using a special purpose computer, a programmed microprocessor or microcontroller and peripheral integrated circuit elements, an ASIC or other integrated circuit, a digital signal processor, a hardwired electronic or logic circuit such as a discrete element circuit, a programmable logic device such as a PLD, PLA, FPGA or PAL, or the like. In general, any device, capable of implementing a finite state machine that is in turn capable of implementing the flowcharts shown in Figs. 2-5 can be used to implement the system for predicting the usage of the a web site 100.

[0057] Each of the circuits 25-80 of the system for predicting the usage of the a web site 100 outlined above can be implemented as portions of a suitably programmed general purpose computer. Alternatively, circuits 25-80 of the system for predicting the usage of the a web site 100 outlined above can be implemented as

physically distinct hardware circuits within an ASIC, or using a FPGA, a PDL, a PLA or a PAL, or using discrete logic elements or discrete circuit elements. The particular form each of the circuits 25-80 of the system for predicting the usage of the a web site 100 outlined above will take is a design choice and will be obvious and predicable to those skilled in the art.

[0058] Moreover, the system for predicting the usage of the a web site 100 and/or each of the various circuits discussed above can each be implemented as software routines, managers or objects executing on a programmed general purpose computer, a special purpose computer, a microprocessor or the like. In this case, the system for predicting the usage of the a web site 100 and/or each of the various circuits discussed above can each be implemented as one or more routines embedded in the communications network, as a resource residing on a server, or the like. The system for predicting the usage of the a web site 100 and the various circuits discussed above can also be implemented by physically incorporating the system for predicting the usage of the a web site 100 into a software and/or hardware system, such as the hardware and software systems of a web server.

[0059] As shown in Fig. 1, the memory, 20 can be implemented using any appropriate combination of alterable, volatile or non-volatile memory or non-alterable, or fixed, memory. The alterable memory, whether volatile or non-volatile, can be implemented using any one or more of static or dynamic RAM, a floppy disk and disk drive, a write-able or rewrite-able optical disk and disk drive, a hard drive, flash memory or the like. Similarly, the non-alterable or fixed memory can be implemented using any one or more of ROM, PROM, EPROM, EEPROM, an optical ROM disk, such as a CD-ROM or DVD-ROM disk, and disk drive or the like.

[0060] The communication links 110 shown in Fig. 1 can each be any known or later developed device or system for connecting a communication device to the system for predicting the usage of the a web site 100, including a direct cable connection, a connection over a wide area network or a local area network, a connection over an intranet, a connection over the Internet, or a connection over any other distributed processing network or system. In general, the communication link 110 can be any known or later developed connection system or structure usable to connect devices and facilitate communication

[0061] Further, it should be appreciated that the communication link 110 can be a wired or wireless link to a network. The network can be a local area network, a wide area network, an intranet, the Internet, or any other distributed processing and storage network.

5 [0062] While this invention has been described in conjunction with the exemplary embodiments outlines above, it is evident that many alternatives, modifications and variations will be apparent to those skilled in the art. Accordingly, the exemplary embodiments of the invention, as set forth above, are intended to be illustrative, not limiting. Various changes may be made without departing from the  
10 spirit and scope of the invention.

11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044  
1045  
1046  
1047  
1048  
1049  
1050  
1051  
1052  
1053  
1054  
1055  
1056  
1057  
1058  
1059  
1060  
1061  
1062  
1063  
1064  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102  
1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
1911  
1912  
1913  
1914  
1915  
1916  
1917  
1918  
1919  
1920  
1921  
1922  
1923  
1924  
1925  
1926  
1927  
1928  
1929  
1930  
1931  
1932  
1933  
1934  
1935  
1936  
1937  
1938  
1939  
1940  
1941  
1942  
1943  
1944  
1945  
1946  
1947  
1948  
1949  
1950  
1951  
1952  
1953  
1954  
1955  
1956  
1957  
1958  
1959  
1960  
1961  
1962  
1963  
1964  
1965  
1966  
1967  
1968  
1969  
1970  
1971  
1972  
1973  
1974  
1975  
1976  
1977  
1978  
1979  
1980  
1981  
1982  
1983  
1984  
1985  
1986  
1987  
1988  
1989  
1990  
1991  
1992  
1993  
1994  
1995  
1996  
1997  
1998  
1999  
2000  
2001  
2002  
2003  
2004  
2005  
2006  
2007  
2008  
2009  
2010  
2011  
2012  
2013  
2014  
2015  
2016  
2017  
2018  
2019  
2020  
2021  
2022  
2023  
2024  
2025  
2026  
2027  
2028  
2029  
2030  
2031  
2032  
2033  
2034  
2035  
2036  
2037  
2038  
2039  
2040  
2041  
2042  
2043  
2044  
2045  
2046  
2047  
2048  
2049  
2050  
2051  
2052  
2053  
2054  
2055  
2056  
2057  
2058  
2059  
2060  
2061  
2062  
2063  
2064  
2065  
2066  
2067  
2068  
2069  
2070  
2071  
2072  
2073  
2074  
2075  
2076  
2077  
2078  
2079  
2080  
2081  
2082  
2083  
2084  
2085  
2086  
2087  
2088  
2089  
2090  
2091  
2092  
2093  
2094  
2095  
2096  
2097  
2098  
2099  
2100  
2101  
2102  
2103  
2104  
2105  
2106  
2107  
2108  
2109  
2110  
2111  
2112  
2113  
2114  
2115  
2116  
2117  
2118  
2119  
2120  
2121  
2122  
2123  
2124  
2125  
2126  
2127  
2128  
2129  
2130  
2131  
2132  
2133  
2134  
2135  
2136  
2137  
2138  
2139  
2140  
2141  
2142  
2143  
2144  
2145  
2146  
2147  
2148  
2149  
2150  
2151  
2152  
2153  
2154  
2155  
2156  
2157  
2158  
2159  
2160  
2161  
2162  
2163  
2164  
2165  
2166  
2167  
2168  
2169  
2170  
2171  
2172  
2173  
2174  
2175  
2176  
2177  
2178  
2179  
2180  
2181  
2182  
2183  
2184  
2185  
2186  
2187  
2188  
2189  
2190  
2191  
2192  
2193  
2194  
2195  
2196  
2197  
2198  
2199  
2200  
2201  
2202  
2203  
2204  
2205  
2206  
2207  
2208  
2209  
2